

=> D4HIS

(FILE 'HOME' ENTERED AT 12:34:41 ON 02 AUG 2003)

FILE 'REGISTRY' ENTERED AT 12:34:50 ON 02 AUG 2003

L1 0 S MITKPGGITXTE/SQSP
L2 0 S VKXTGIP1/SQSP
L3 88 S ZPDIII/SQSP
L4 88 DUP REM L3 (0 DUPLICATES REMOVED)
L5 0 S ZPDIIIXXP/SQSP
L6 0 S ZPDIIIXXP/SQSP
L7 3 S EHQPDI11/SQSP
L8 5 S QVvvvCGKN/SQSP

FILE 'REGISTRY' ENTERED AT 12:48:05 ON 02 AUG 2003

SET TERMSET E#
DEL SEL Y
SEL L8 5 RN
L9 1 S E1/RN
SET TERMSET LOGIN

FILE 'CPLUS' ENTERED AT 12:48:09 ON 02 AUG 2003

L10 5 S L9
S DCMITKPG/SQSP

FILE 'REGISTRY' ENTERED AT 12:51:40 ON 02 AUG 2003

L11 3 S DCMITKPG/SQSP

FILE 'CPLUS' ENTERED AT 12:51:43 ON 02 AUG 2003

L12 5 S L11

=> d 17 SQIDE 1-3

L7 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN
RN 459580-03-1 REGISTRY
CN GenBank CAB14110 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank CAB14110 (Translated from: GenBank Z99115)
FS PROTEIN SEQUENCE
SQL 382

SEQ 1 MNTNKRVLIL TANYGNHVG VAKTLYEQCV RLGFQHVTVS NLYQESNPIV
51 SEVTQYLYLK SFSIGKQFYR LFYYGVDKIY NKRKFNIYFK MGNKRLGELV
101 DEHQPDIIIN TFPMIVVPEY RRRTGRVIPT FNVMDFCLH KIWVHENVDK
=====

151 YYVATDYVKE KLLEIGTHPS NVKITGIPIR PQFEESMPVG PIYKKYNLSP
201 NKKVLLIMAG AHGVLKNVKE LCENLVKDDQ VQVVVVCGKN TALKESLSAL
251 EAENGDKLKV LGYVERIDEL FRITDCMITK PGGITLTEAT AIGVPVILYK
301 PVPGQEKENA NFFEDRGAAI VVNRHEEILE SVTSLLADED TLHRMKKNIK
351 DLHLANSSEV ILEDILKESE MMTAKQKAKV LS

HITS AT: 102-109

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA

L7 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN
RN 459580-02-0 REGISTRY
CN GenBank AAA96624 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAA96624 (Translated from: GenBank L77246)
FS PROTEIN SEQUENCE
SQL 382

SEQ 1 MNTNKRVLIL TANYGNHVG VAKTLYEQCV RLGFQHVTVS NLYQESNPIV
51 SEVTQYLYLK SFSIGKQFYR LFYYGVDKIY NKRKFNIYFK MGNKRLGELV
101 DEHQPDIIIN TFPMIVVPEY RRRTGRVIPT FNVMDFCLH KIWVHENVDK
=====

151 YYVATDYVKE KLLEIGTHPS NVKITGIPIR PQFEESMPVG PIYKKYNLSP
201 NKKVLLIMAG AHGVLKNVKE LCENLVKDDQ VQVVVVCGKN TALKESLSAL
251 EAENGDKLKV LGYVERIDEL FRITDCMITK PGGITLTEAT AIGVPVILYK
301 PVPGQEKENA NFFEDRGAAI VVNRHEEILE SVTSLLADED TLHRMKKNIK
351 DLHLANSSEV ILEDILKESE MMTAKQKAKV LS

HITS AT: 102-109

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA

L7 ANSWER 3 OF 3 REGISTRY COPYRIGHT 2003 ACS on STN
RN 185159-01-7 REGISTRY
CN Glycosyltransferase (Bacillus subtilis strain 168 clone YAC10-9 gene ypfP)
(9CI) (CA INDEX NAME)
OTHER NAMES:
CN 13: PN: US6251647 SEQID: 12 unclaimed protein
CN Protein (Bacillus subtilis gene ypfP)
CN Protein (Bacillus subtilis strain 168 clone YAC10-9 gene ypfP)
FS PROTEIN SEQUENCE
SQL 382

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference

=====+=====+
Not Given | US6251647
| unclaimed
| SEQID 12

SEQ 1 MNTNKRVLIL TANYGNHVG VAKTLYEQCV RLGFQHVTVS NLYQESNPIV
51 SEVTQYLYLK SFSIGKQFYR LFYYGVDKIY NKRKFNIYFK MGNKRLGELV
101 DEHQPDIIIN TFPMIVVPEY RRRTGRVIPT FNVMDFCLH KIWVHENVDK
=====+
151 YYVATDYVKE KLLEIGTHPS NVKITGIPIR PQFEESMPVG PIYKKYNLSP
201 NKKVLLIMAG AHGVLKVNKE LCENLVKDDQ VQVVVVCGKN TALKESLSAL
251 EAENGDKLKVG LGYVERIDEL FRITDCMITK PGGITLTEAT AIGVPVILYK
301 PVPGQEKENA NFFEDRGAAI VVNRHEEILE SVTSSLADED TLHRMKKNIK
351 DLHLANSSEV ILEDILKESE MMTAKQKAKV LS

HITS AT: 102-109

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

5 REFERENCES IN FILE CA (1947 TO DATE)

5 REFERENCES IN FILE CAPLUS (1947 TO DATE)

=> D L8 SQIDE 1-5

L8 ANSWER 1 OF 5 REGISTRY COPYRIGHT 2003 ACS on STN
RN 516397-22-1 REGISTRY
CN Protein (Bacillus anthracis strain Ames gene BA0511) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAP24533
CN GenBank AAP24533 (Translated from: GenBank AE017025)
FS PROTEIN SEQUENCE
SQL 388

SEQ 1 MIKNPKVLIL TAHYGNHVQ VAKTLEQTFR QKGKDVIVC DLFGESHPVI
51 TDITKYLYLK SYTIGKELYR LFYYGVEKIY DKKIASWYAN FGRKRLKLLL
101 QAEKPDIVIN TFPIIAVPEL KKQQTGISIPV YNVLTDFCVH KIWIHREVDR
151 YFVATDHVKK VMVDIGVPAE QIVETGIPIR SSFELKINPD IIYNKYQLCK
201 NKKILLIVAG AHGVLGSVKE LCQSFMSVPD LQVVVVCGKN EALKQDLVGV
=====

251 QETNPDALKV FGYVENIDEL FRVTSCMITK PGGITLSEAA ALQVPVILYK
301 PVPGQENENA MYFERKGAAV VIRDDSEVFA KTEALLQDDM KLLQMKEAMK
351 SIYRPEPADH IVDTILAENH VEPNHIPIKS PALAQSF

HITS AT: 232-240

MF Unspecified

CI MAN

SR GenBank

LC STN Files: CA, CAPLUS

1 REFERENCES IN FILE CA (1947 TO DATE)
1 REFERENCES IN FILE CAPLUS (1947 TO DATE)

L8 ANSWER 2 OF 5 REGISTRY COPYRIGHT 2003 ACS on STN
RN 508956-41-0 REGISTRY
CN 1,2-Diacylglycerol 3-glucosyltransferase (Bacillus cereus strain ATCC 14579 gene BC0493) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAP07531
CN GenBank AAP07531 (Translated from: GenBank AE016999)
FS PROTEIN SEQUENCE
SQL 388

SEQ 1 MIKNPKVLIL TAHYGNHVQ VAKTLEQTFR QKGIEDVIVC DLFGESHPFI
51 TDITKYLYLK SYTIGKELYR LFYYGVEKIY DKKIASWYAN FGRKRLKTLL
101 QVEKPDIVIN TFPIIAVPEL KKQQTGISIPV YNVLTDFCVH KIWIHREVDR
151 YFVATDHVKE LMVDIGVPAE QIVETGIPIR SSFELKVNPE IIYTKYQLCK
201 NKKILLIVAG AHGVLGNVKE LCQSFMSVPN LQVVVVCGKN EALKQDLSSL
=====

251 QKQNSDALKV FGYVENIDEL FRVTSCMITK PGGITLSEAA ALQVPVILYK
301 PVPGQENENA MYFERKGAAV VIRDDSEVFA KTEALLQDDV KLLQMKEAMK
351 SIYLPEPAGH IVDAILAENH AEPRHIPIKS PALAQSF

HITS AT: 232-240

MF Unspecified

CI MAN

SR GenBank

LC STN Files: CA, CAPLUS, TOXCENTER

1 REFERENCES IN FILE CA (1947 TO DATE)
1 REFERENCES IN FILE CAPLUS (1947 TO DATE)

L8 ANSWER 3 OF 5 REGISTRY COPYRIGHT 2003 ACS on STN
RN 459580-03-1 REGISTRY
CN GenBank CAB14110 (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank CAB14110 (Translated from: GenBank Z99115)
FS PROTEIN SEQUENCE
SQL 382

SEQ 1 MNTNKRLVIL TANYGNHVG VAKTLYEQCV RLGFQHVTVS NLYQESNPIV
51 SEVTQYLYLK SFSIGKQFYR LFYYGVDKIY NKRKFNIYFK MGNKRLGELV
101 DEHQPDIIIN TFPMIIVVPEY RRRRTGRVIPT FNVMDFCLH KIWHENVDK
151 YYVATDYVKE KLLEIGTHPS NVKITGIPIR PQFEESMPVG PIYKKYNLSP
201 NKKVLLIMAG AHGVLKNVKE LCENLVKDDQ VQVVVVCGKN TALKESLSAL
=====

251 EAENGDKLKV LGYVERIDEL FRITDCMITK PGGITLTEAT AIGVPVILYK
301 PVPGQEKENA NFFEDRGAAI VVNRHEEILE SVTSLLADED TLHRMKNNIK
351 DLHLANSSEV ILEDILKESE MMTAKQKAKV LS

HITS AT: 232-240

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA

L8 ANSWER 4 OF 5 REGISTRY COPYRIGHT 2003 ACS on STN
RN 459580-02-0 REGISTRY
CN GenBank AAA96624 (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAA96624 (Translated from: GenBank L77246)
FS PROTEIN SEQUENCE
SQL 382

SEQ 1 MNTNKRLVIL TANYGNHVG VAKTLYEQCV RLGFQHVTVS NLYQESNPIV
51 SEVTQYLYLK SFSIGKQFYR LFYYGVDKIY NKRKFNIYFK MGNKRLGELV
101 DEHQPDIIIN TFPMIIVVPEY RRRRTGRVIPT FNVMDFCLH KIWHENVDK
151 YYVATDYVKE KLLEIGTHPS NVKITGIPIR PQFEESMPVG PIYKKYNLSP
201 NKKVLLIMAG AHGVLKNVKE LCENLVKDDQ VQVVVVCGKN TALKESLSAL
=====

251 EAENGDKLKV LGYVERIDEL FRITDCMITK PGGITLTEAT AIGVPVILYK
301 PVPGQEKENA NFFEDRGAAI VVNRHEEILE SVTSLLADED TLHRMKNNIK
351 DLHLANSSEV ILEDILKESE MMTAKQKAKV LS

HITS AT: 232-240

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA

L8 ANSWER 5 OF 5 REGISTRY COPYRIGHT 2003 ACS on STN
RN 185159-01-7 REGISTRY
CN Glycosyltransferase (Bacillus subtilis strain 168 clone YAC10-9 gene ypfP)
(9CI) (CA INDEX NAME)

OTHER NAMES:

CN 13: PN: US6251647 SEQID: 12 unclaimed protein
CN Protein (Bacillus subtilis gene ypfP)
CN Protein (Bacillus subtilis strain 168 clone YAC10-9 gene ypfP)
FS PROTEIN SEQUENCE
SQL 382

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
=====	
Not Given	US6251647
	unclaimed
	SEQID 12

SEQ 1 MNTNKRLVIL TANYGNHVG VAKTLYEQCV RLGFQHVTVS NLYQESNPIV
51 SEVTQYLYLK SFSIGKQFYR LFYYGVDKIY NKRKFNIYFK MGNKRLGELV
101 DEHQPDIIIN TFPMIIVVPEY RRRRTGRVIPT FNVMDFCLH KIWHENVDK
151 YYVATDYVKE KLLEIGTHPS NVKITGIPIR PQFEESMPVG PIYKKYNLSP

201 NKKVLLIMAG AHGVLKNVKE LCENLVKDDQ VQVVVVCGKN TALKESLSAL

=====

251 EAENGDKLKV LGYVERIDEL FRITDCMITK PGGITLTEAT AIGVPVILYK

301 PVPGQEKENA NFFEDRGAAI VVNRHEEILE SVTSLLADED TLHRMKKNIK

351 DLHLANSSEV ILEDILKESE MMTAKQKAKV LS

HITS AT: 232-240

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

5 REFERENCES IN FILE CA (1947 TO DATE)

5 REFERENCES IN FILE CAPLUS (1947 TO DATE)

=> DIS L10 1 IBIB ABS

THE ESTIMATED COST FOR THIS REQUEST IS 2.42 U.S. DOLLARS
DO YOU WANT TO CONTINUE WITH THIS REQUEST? (Y)/N:Y

L10 ANSWER 1 OF 5 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 2001:468188 CAPLUS
DOCUMENT NUMBER: 135:73946
TITLE: Auxiliary genes and proteins of methicillin resistant *Staphylococcus aureus* and the development of antibiotics against these strains
INVENTOR(S): De Lencastre, Herminia; Tomasz, Alexander
PATENT ASSIGNEE(S): Rockefeller University, USA
SOURCE: U.S., 76 pp., Cont.-in-part of U.S. Ser. No. 961,595, abandoned.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 3
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6251647	B1	20010626	US 1997-984618	19971203
WO 9516039	A1	19950615	WO 1994-US13952	19941206
W: AU, CA, JP, US				
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 6063613	A	20000516	US 1995-403918	19950315
PRIORITY APPLN. INFO.:			US 1993-163053	B2 19931206
			WO 1994-US13952	A2 19941206
			US 1995-403918	A2 19950315
			US 1997-961595	B2 19971031

AB The present invention is directed to the identification of mutant strains of methicillin resistant bacteria, in particular methicillin resistant *Staphylococcus aureus* (MRSA), to identify the characteristics of such bacteria and develop drugs that can reverse, inhibit, or reduce bacterial resistance to β -lactam antibiotics, e.g., methicillin. The invention particularly relates to identification of a novel mutant strain of methicillin resistant *S. aureus* that manifests a unique phenotype. Accordingly, the invention provides for methods of treatment and corresponding pharmaceutical compns. for treating bacterial, particularly staphylococcal, infections. Satn. transposon mutagenesis of the chromosome of MRSA strain COL with Tn551 identified 70 events leading to lowered methicillin resistance, 58 of which were at new insertion sites. These genes may therefore be needed in addn. to the *mecA* gene to maximize methicillin resistance. Some of the mutations affect the compn. of the cell wall. Specifically, mutations in the *murE* gene affecting methicillin resistance are identified.

REFERENCE COUNT: 48 THERE ARE 48 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=> DIS L10 2 IBIB ABS

THE ESTIMATED COST FOR THIS REQUEST IS 2.42 U.S. DOLLARS
DO YOU WANT TO CONTINUE WITH THIS REQUEST? (Y)/N:Y

L10 ANSWER 2 OF 5 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 1999:626343 CAPLUS
DOCUMENT NUMBER: 131:254319
TITLE: Processive glycosyltransferases of *Bacillus* and *Staphylococcus* and their use in glycolipid synthesis
INVENTOR(S): Wolter, Frank P.; Jorasch, Petra; Heinz, Ernst; Zahringer, Ulrich
PATENT ASSIGNEE(S): GVS Gesellschaft fur Erwerb und Verwertung Landwirtschaftlicher Pflanzensort, Germany;

SOURCE: Forschungszentrum Borstel
PCT Int. Appl., 37 pp.
CODEN: PIXXD2

DOCUMENT TYPE: Patent
LANGUAGE: German
FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9949052	A2	19990930	WO 1999-DE857	19990325
WO 9949052	A3	20000302		
W: AU, CA, CZ, HU, PL, SI, US				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
DE 19819958	A1	19990930	DE 1998-19819958	19980505
CA 2329898	AA	19990930	CA 1999-2329898	19990325
AU 9941301	A1	19991018	AU 1999-41301	19990325
AU 759792	B2	20030501		
EP 1066388	A2	20010110	EP 1999-924670	19990325
R: AT, BE, CH, DE, DK, FR, GB, LI, NL, SE, IE				
PRIORITY APPLN. INFO.:			DE 1998-19813017 A	19980325
			DE 1998-19819958 A	19980505
			WO 1999-DE857 W	19990325

AB The title enzymes and their use are disclosed. Thus, the ypfP gene of *B. subtilis* and of *S. aureus* were expressed in *Escherichia coli*. Both enzymes utilized UDP-glucose, and catalyzed addn. of up to 4 glucosyl moieties in β -(1.fwdarw.6) linkage to the substrates. The *Bacillus* enzyme used diacylglycerol, monoglycosyl diacylglycerol, diglycosyl diacylglycerol and alkyl- α ./ β -D-glucopyranosides as acceptor. The *Staphylococcus* enzyme could also use sterols and sterolglucosids as acceptors. Two novel phosphoglycolipids were identified.

=> DIS L10 3 IBIB ABS
THE ESTIMATED COST FOR THIS REQUEST IS 2.42 U.S. DOLLARS
DO YOU WANT TO CONTINUE WITH THIS REQUEST? (Y)/N:Y

L10 ANSWER 3 OF 5 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 1997:748948 CAPLUS
DOCUMENT NUMBER: 128:150233
TITLE: The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*
AUTHOR(S): Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A. M.; Alloni, G.; Azevedo, V.; Bertero, M. G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Borriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S. C.; Bron, S.; Brouillet, S.; Bruschi, C. V.; Caldwell, B.; Capuano, V.; Carter, N. M.; Choi, S.-K.; Codani, J.-J.; Connerton, I. F.; Cummings, N. J.; Daniel, R. A.; Denizot, F.; Devine, K. M.; Dusterhoff, A.; Ehrlich, S. D.; Emmerson, P. T.; Entian, K. D.; Errington, J.; Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.-Y.; Glaser, P.; Goffeau, A.; Golightly, E. J.; Grandi, G.; Guiseppi, G.; Guy, B. J.; Haga, K.; et al.
CORPORATE SOURCE: Unite de Biochemie Microbienne, Inst. Pasteur, Paris, 75724, Fr.
SOURCE: Nature (London) (1997), 390(6657), 249-256
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Macmillan Magazines
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Bacillus subtilis is the best-characterized member of the gram-pos. bacteria. Its genome of 4,214,810 base pairs comprises 4100 protein-coding genes. Of these protein-coding genes, 53% are represented once, while a quarter of the genome corresponds to several gene families that have been greatly expanded by gene duplication, the largest family contg. 77 putative ATP-binding transport proteins. In addn., a large proportion of the genetic capacity is devoted to the utilization of a variety of carbon sources, including many plant-derived mols. The identification of 5 signal peptidase genes, as well as several genes for components of the secretion app., is important given the capacity of Bacillus strains to secrete large amts. of industrially important enzymes. Many of the genes are involved in the synthesis of secondary metabolites, including antibiotics, that are more typically assocd. with Streptomyces species. The genome contains >10 prophages or remnants of prophages, indicating that bacteriophage infection has played an important evolutionary role in horizontal gene transfer, in particular in the propagation of bacterial pathogenesis.

REFERENCE COUNT: 49 THERE ARE 49 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=> DIS L10 4 IBIB ABS
THE ESTIMATED COST FOR THIS REQUEST IS 2.42 U.S. DOLLARS
DO YOU WANT TO CONTINUE WITH THIS REQUEST? (Y)/N:Y

L10 ANSWER 4 OF 5 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 1997:497492 CAPLUS
DOCUMENT NUMBER: 127:201862
TITLE: A Bacillus subtilis gene encoding a protein similar to nucleotide sugar transferases influences cell shape and viability
AUTHOR(S): Price, Kirsten D.; Roels, Steven; Losick, Richard
CORPORATE SOURCE: Dep. Mol. Cell. Biol., Biological Laboratories, Harvard Univ., Cambridge, MA, 02138, USA
SOURCE: Journal of Bacteriology (1997), 179(15), 4959-4961
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Bacillus subtilis gene ypfP, which is located at 196.degree. on the genetic map, shows similarity to both the monogalactosyldiacylglycerol synthase gene of Cucumis sativus, which encodes a galactosyltransferase, and the murG genes of B. subtilis, Escherichia coli, Haemophilus influenzae, and Synechocystis sp. strain PCC6803, which encode N-acetylglucosaminyltransferases involved in peptidoglycan biosynthesis. Cells contg. a null mutation of ypfP are shorter and rounder than wild-type cells during growth in Luria-Bertani medium and glucose minimal medium. In addn., the mutant cells preferentially undergo lysis when grown on solid Luria-Bertani medium.

=> DIS L10 5 IBIB ABS
THE ESTIMATED COST FOR THIS REQUEST IS 2.42 U.S. DOLLARS
DO YOU WANT TO CONTINUE WITH THIS REQUEST? (Y)/N:Y

L10 ANSWER 5 OF 5 CAPLUS COPYRIGHT 2003 ACS on STN
ACCESSION NUMBER: 1996:759403 CAPLUS
DOCUMENT NUMBER: 126:55649
TITLE: Organization of the Bacillus subtilis 168 chromosome between kdg and the attachment site of the SP.beta. prophage: use of long accurate PCR and yeast artificial chromosomes for sequencing
AUTHOR(S): Capuano, Veronique; Galleron, Nathalie; Pujic, Petar; Sorokin, Alexei; Ehrlich, S. Dusko

CORPORATE SOURCE: Lab. Genet. Microbienne, Inst. Natl. Recherche Agronomique, Jouy en Josas, 78352, Fr.
SOURCE: Microbiology (Reading, United Kingdom) (1996), 142(11), 3005-3015
CODEN: MROBEO; ISSN: 1350-0872
PUBLISHER: Society for General Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Within the *Bacillus subtilis* genome sequencing project, the region between *lysA* and *ilvA* was investigated. This report presents the sequence of the last 36 kb of this region, between the *kdg* operon and the attachment site of the *SP.beta.* prophage. A 2-step strategy was used for the sequencing. In the first step, total chromosomal DNA was cloned in phage M13-based vectors and the clones carrying inserts from the target region were identified by hybridization with a cognate yeast artificial chromosome (YAC). Sequencing of the clones allowed establishment of a no. of contigs. In the second step the contigs were mapped by Long Accurate (LA) PCR and the remaining gaps closed by sequencing of the PCR products. The level of sequence inaccuracy due to LA PCR errors appeared to be about 1 in 10,000, which does not affect significantly the final sequence quality. This 2-step strategy is efficient and could be applied to sequencing of longer chromosomal regions. The 36-kb sequence contains 38 coding sequences (CDSs), 19 of which encode unknown proteins. Seven genetic loci already mapped in this region, *xpt*, *metB*, *ilvA*, *ilvD*, *thyB*, *dfrA* and *degR* were identified. Eleven ORFs display significant similarities to known proteins from the data banks, suggesting possible functions for some of the novel genes: *cspD* may encode a cold shock protein; *bcsA*, the first bacterial homolog of chalcone synthase; *exoI*, a 5' to 3' exonuclease, similar to that of DNA polymerase I of *Escherichia coli*; and *bsaA*, a stress-response-assocd. protein. The protein encoded by *yplP* has homol. with the transcriptional *NifA*-like regulators. The arrangement of the genes relative to possible promoters and terminators suggests 19 potential transcription units.

=> D HIS

(FILE 'HOME' ENTERED AT 12:34:41 ON 02 AUG 2003)

FILE 'REGISTRY' ENTERED AT 12:34:50 ON 02 AUG 2003

L1 0 S MITKPGGITXTE/SQSP
L2 0 S VKXTGIP1/SQSP
L3 88 S ZPDIII/SQSP
L4 88 DUP REM L3 (0 DUPLICATES REMOVED)
L5 0 S ZPDIIIXXP/SQSP
L6 0 S ZPDIIIXXXP/SQSP
L7 3 S EHQPDI1/SQSP
L8 5 S QVVVVCGKN/SQSP

FILE 'REGISTRY' ENTERED AT 12:48:05 ON 02 AUG 2003

SET TERMSET E#
DEL SEL Y
SEL L8 5 RN
L9 1 S E1/RN
SET TERMSET LOGIN

FILE 'CAPLUS' ENTERED AT 12:48:09 ON 02 AUG 2003

L10 5 S L9

=> S DCMITKPG/SQSP

REG1stRY INITIATED

Substance data SEARCH and crossover from CAS REGISTRY in progress...
Use DISPLAY HITSTR (or FHITSTR) to directly view retrieved structures.

